Reconocimiento de Patrones

Version 2022-2

Clustering: Mixture of Gaussians

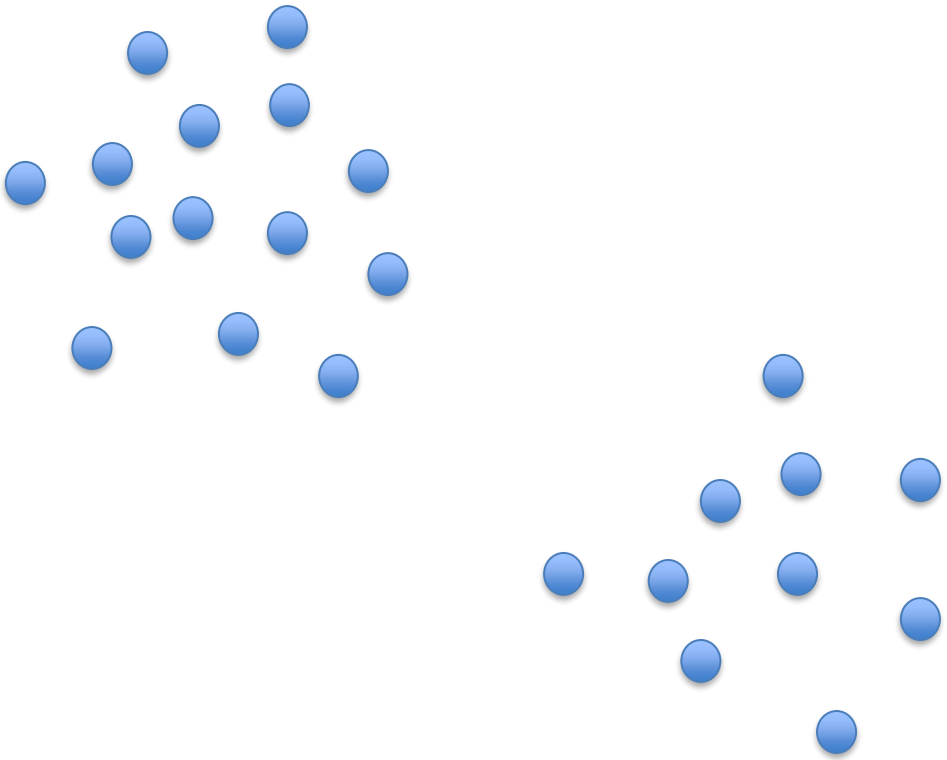
[Capítulo 6]

Dr. José Ramón Iglesias

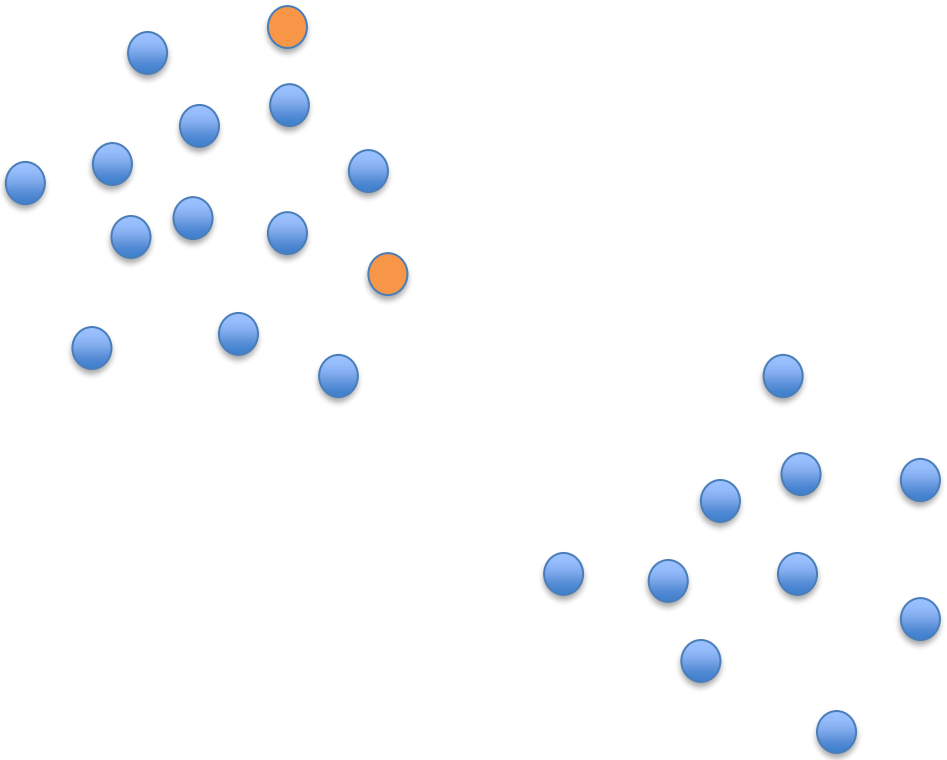
DSP-ASIC BUILDER GROUP
Director Semillero TRIAC
Ingenieria Electronica
Universidad Popular del Cesar

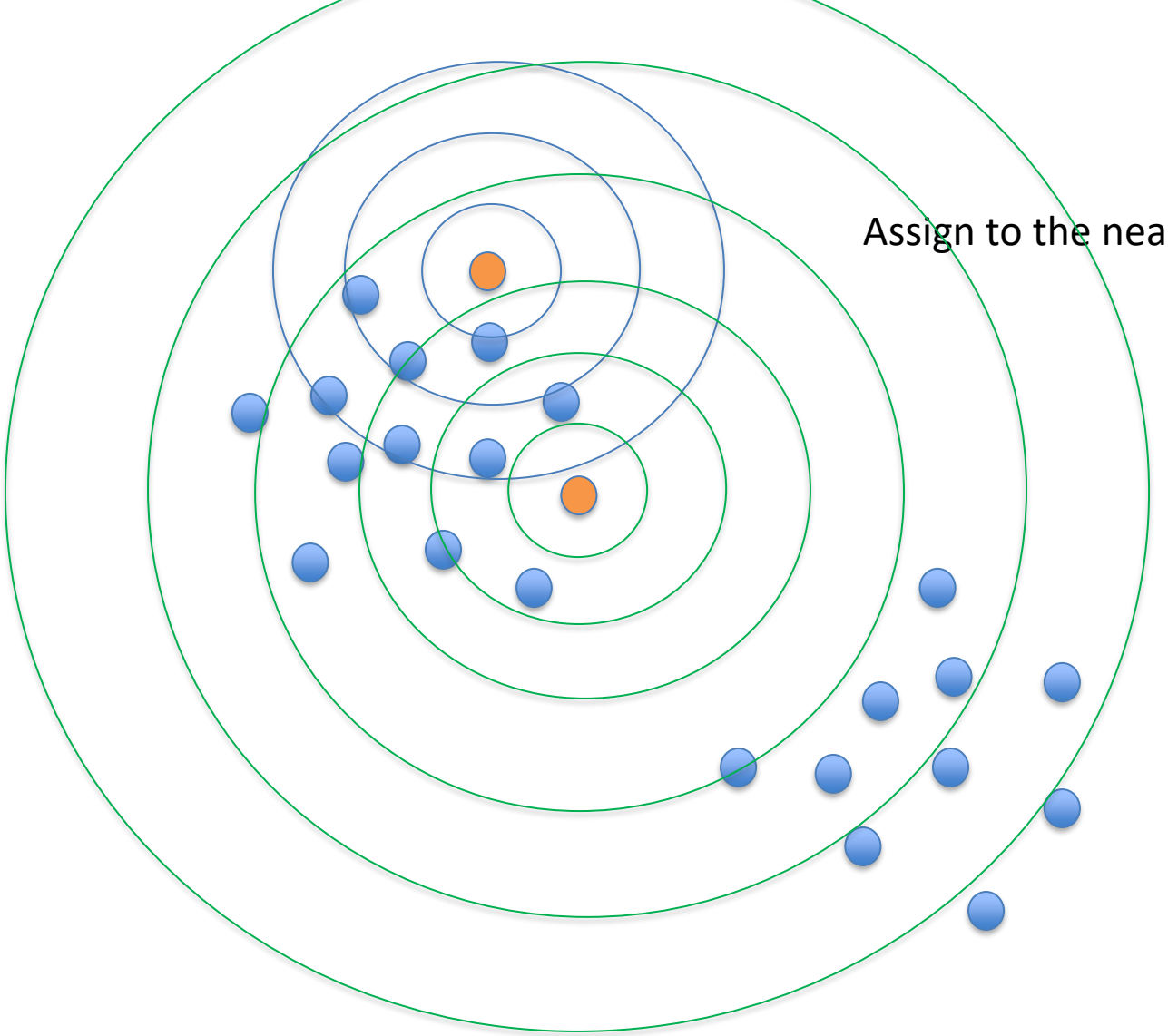
Algorithm:

1. Input Data X = {x_1, x_2, ... x_N} and number of clusters K
2. Centroids {c_1, c_2, ... c_K} = random K points of X
3. Initial values for μ_k, Σ_k
4. For each data point x_i
5. Compute Mahalanobis distance $d_{ik} = d(x_i,c_k)$ i=1,...,N, k=1,...K
6. Assign x_i to the nearest centroid: $y_i = \text{argmin}_j\{d_{ik}\}$
7. Compute for each cluster μ_k, Σ_k ($c^*_k = \mu_k$)
8. if $c^*_k \neq c_k$ then $c_k = c^*_k$ go to step 4
9. Output: {c^*_1, c^*_2, ... c^*_K} and y_i for i=1,...,N

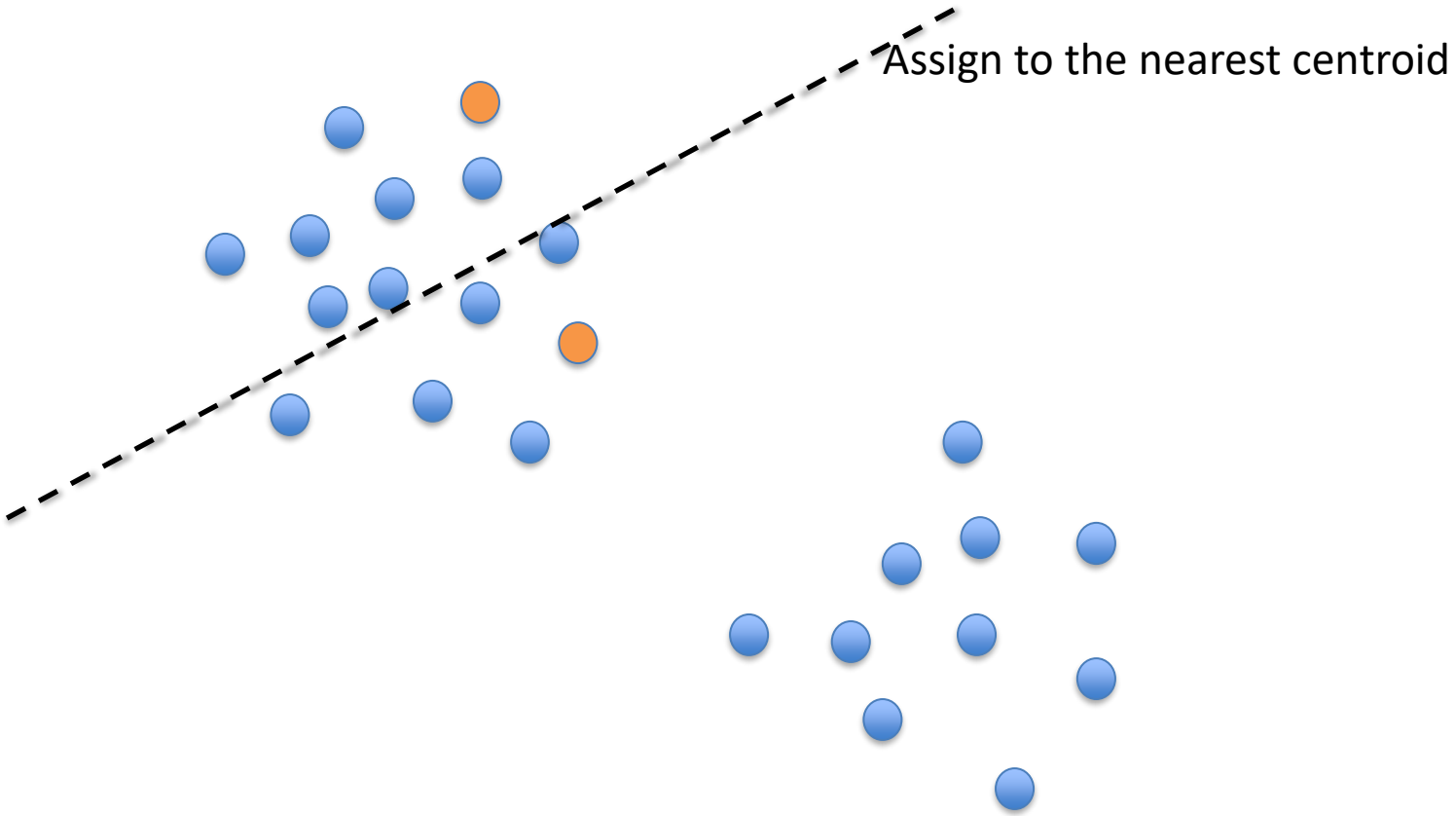


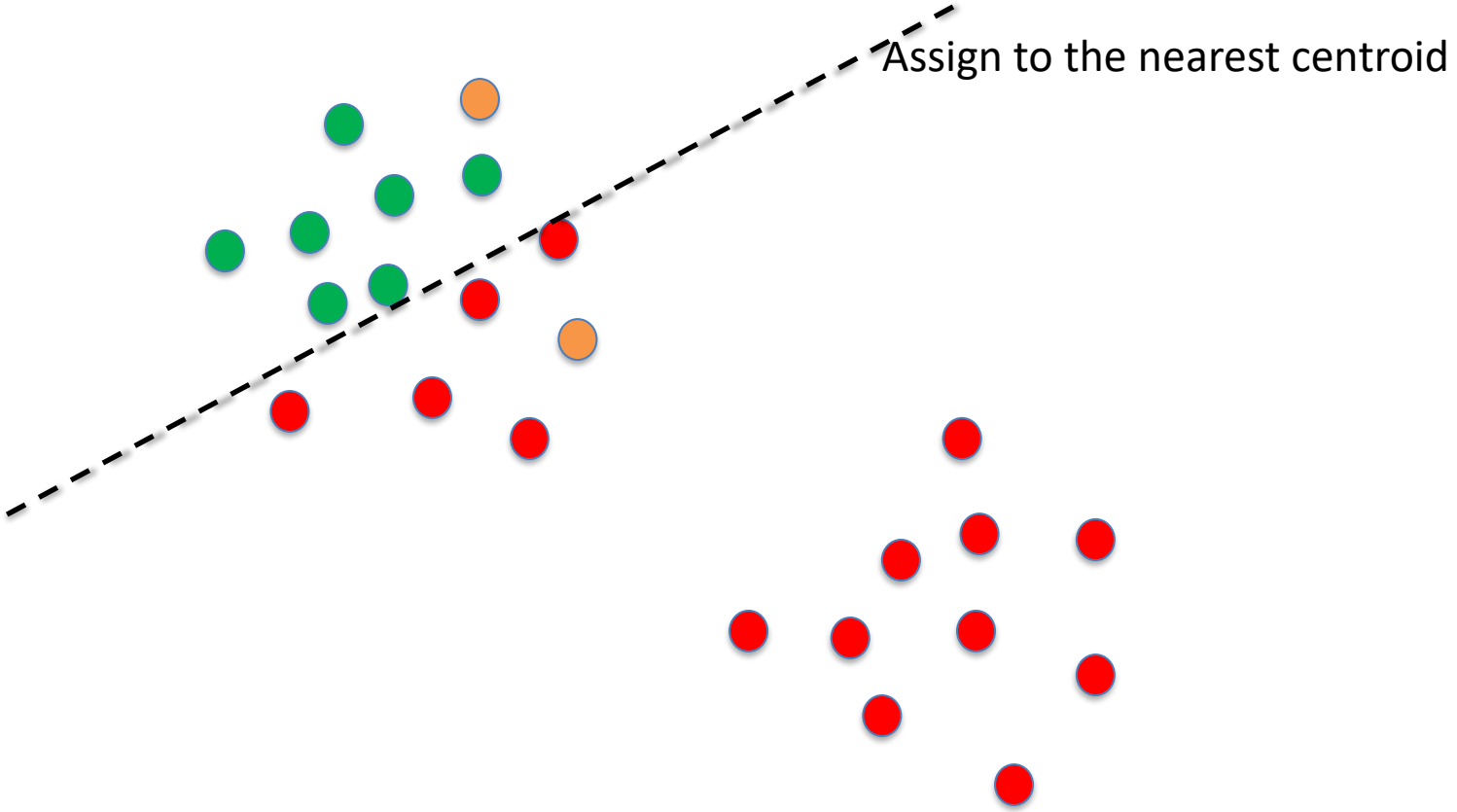
Choose random K=2 points (centroids)

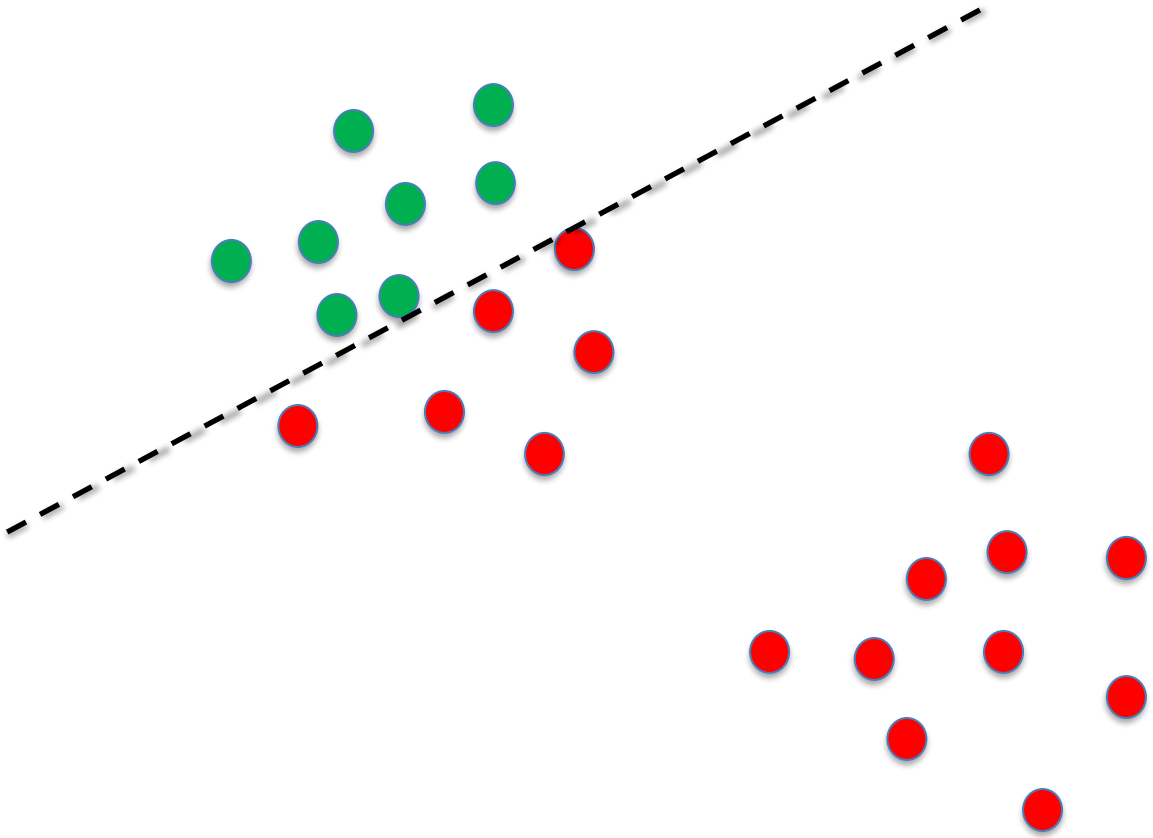




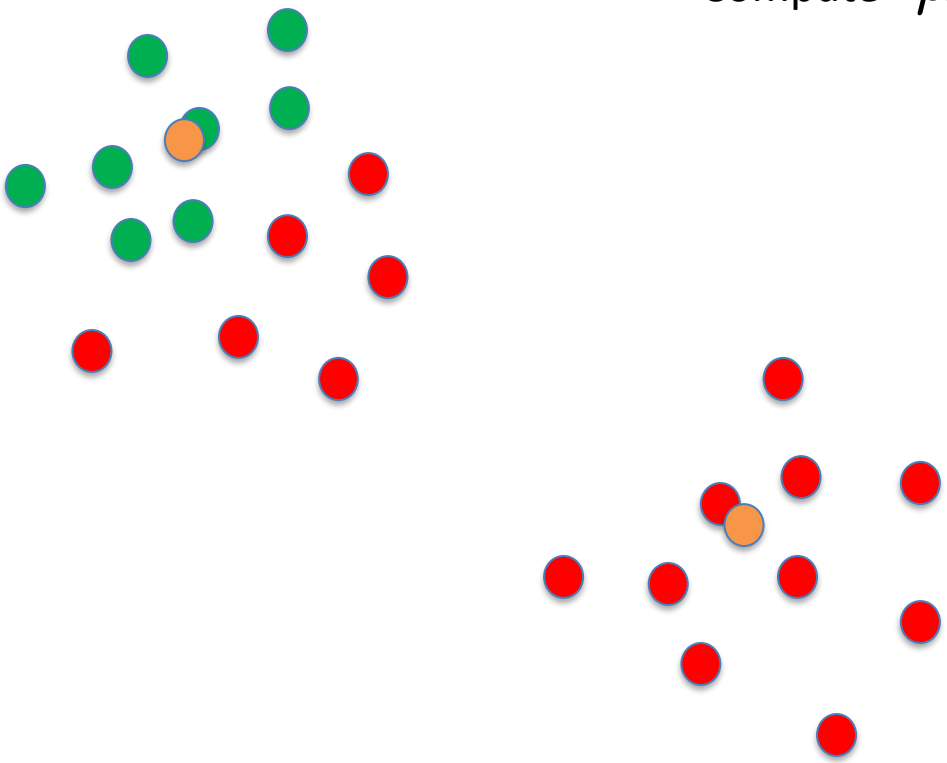
Assign to the nearest centroid

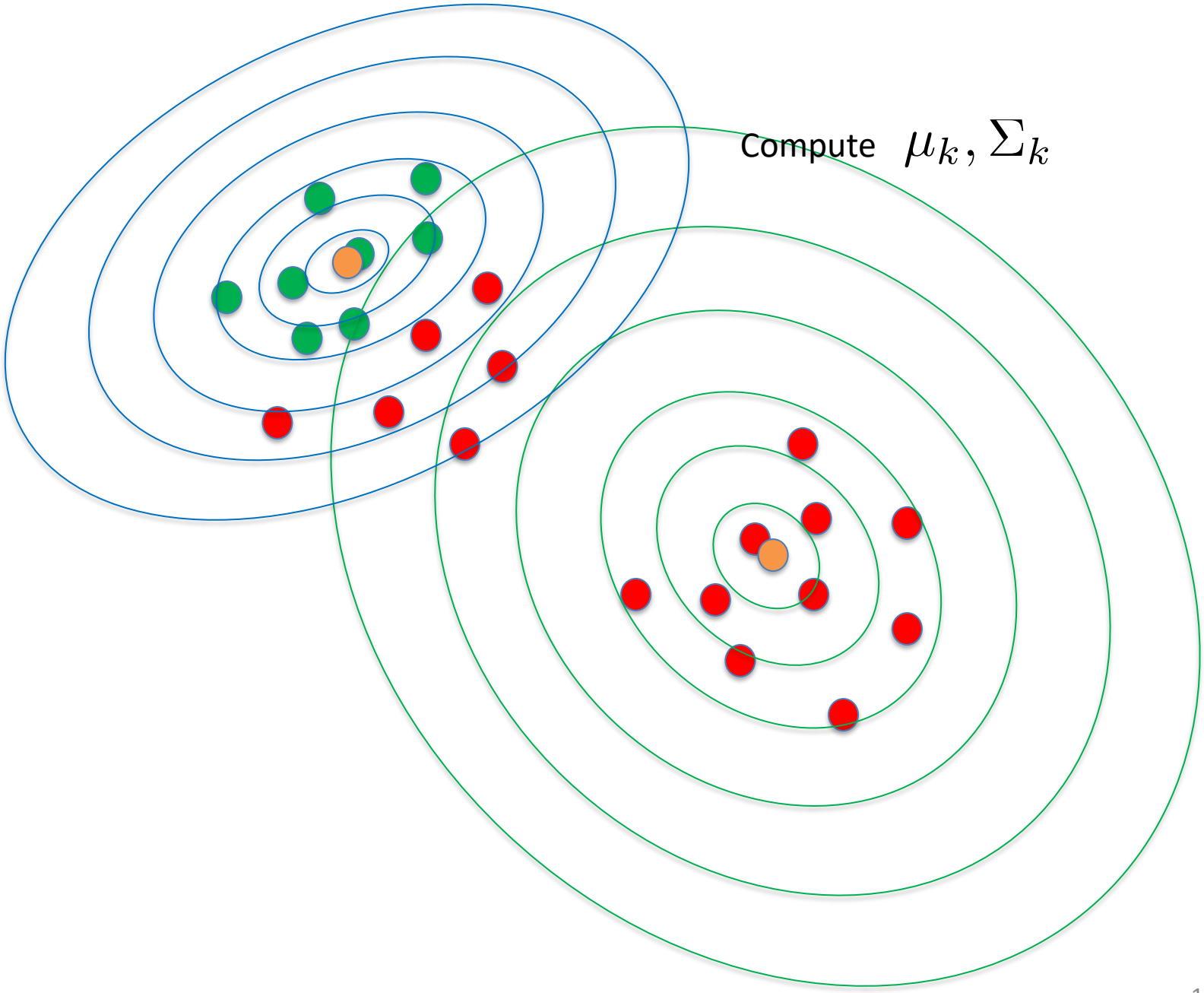

Assign to the nearest centroid

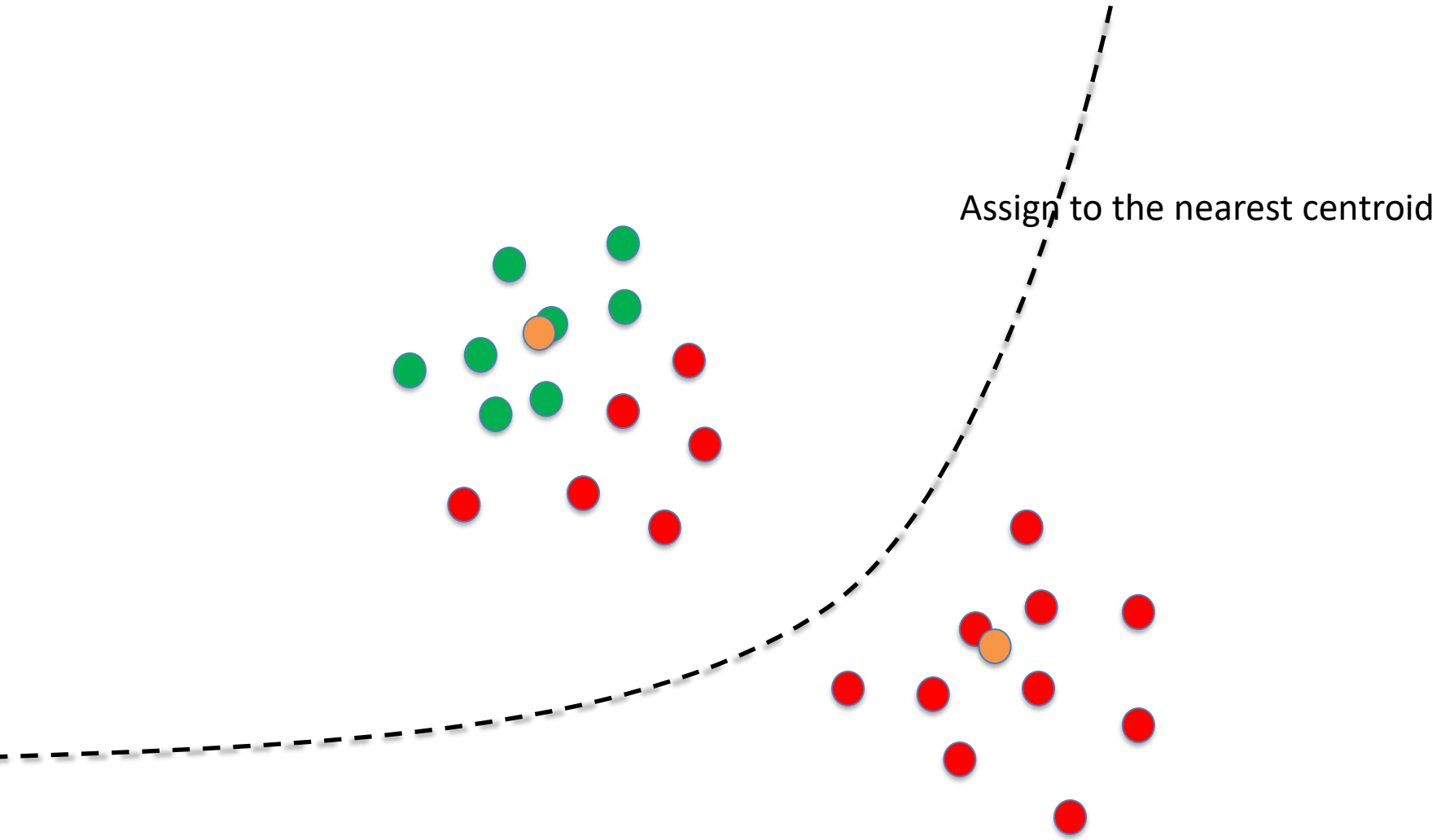

Assign to the nearest centroid

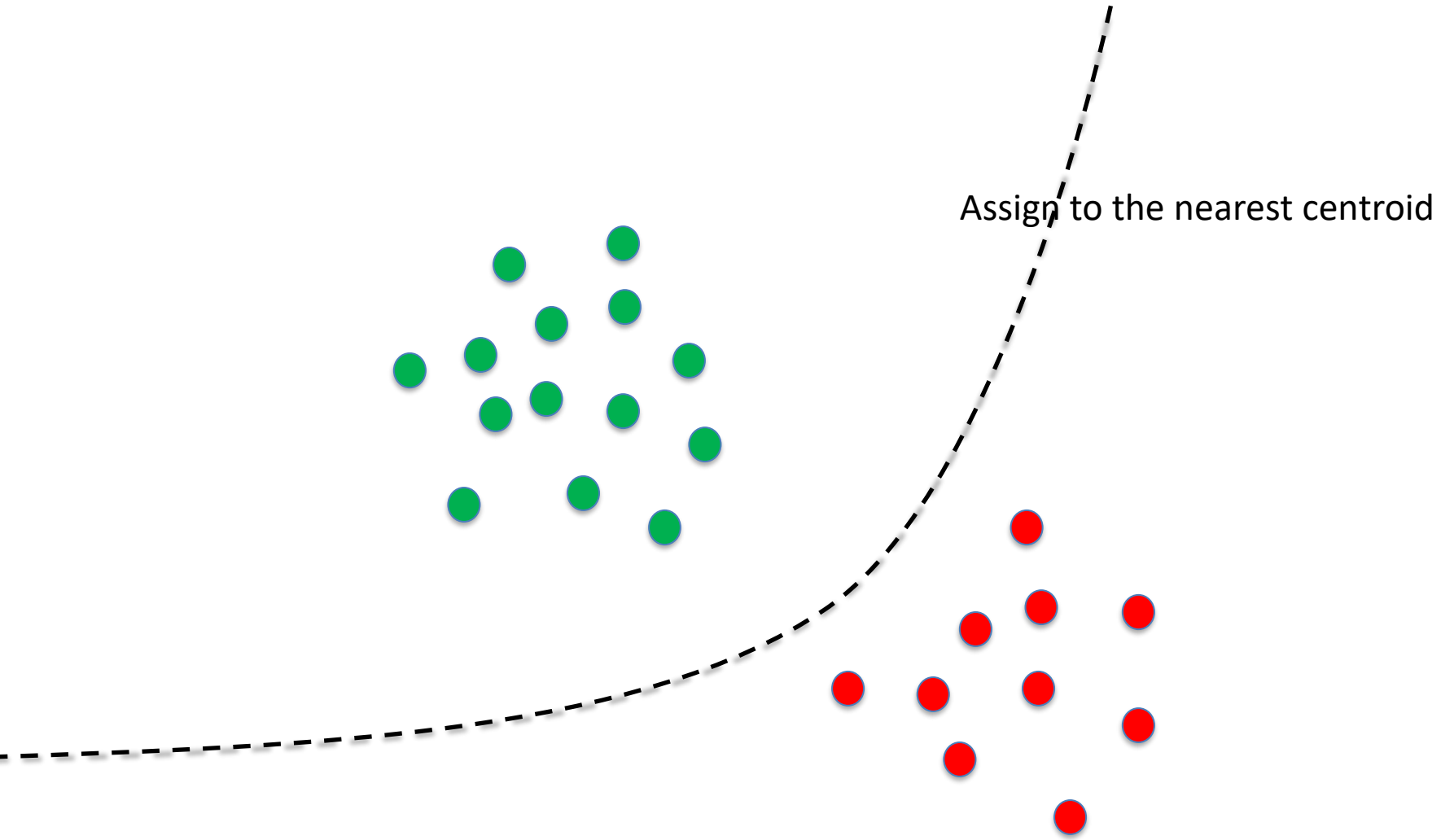


Compute μ_k, Σ_k




Compute μ_k, Σ_k


Assign to the nearest centroid


Assign to the nearest centroid

Repeat until convergence